

Cynthia G. Clopper

Sound change in the individual: Effects of exposure on cross-dialect speech processing

Abstract: Speech perception is highly robust to variation, but familiarity with a particular source of variation can nevertheless lead to significant processing benefits. In the domain of cross-dialect speech perception, familiar local and standard varieties have been shown to facilitate lexical and semantic processing relative to unfamiliar dialects. However, more recent research suggests that individuals with exposure to both a local non-standard variety and a regional or national standard variety exhibit a mix of lexical processing costs and benefits, suggesting that familiarity with multiple different linguistic systems can result in both independent processing benefits for each variety as well as competition among variable multi-dialect representations. In an exemplar model of language processing, this complex pattern of results suggests several loci for sound change within an individual language user. Although the processing benefits associated with the local variety may contribute to long-term maintenance of variation through divergence from a regional or national standard, the processing benefits associated with the standard may contribute to dialect leveling and convergence towards the standard. Competition between these forces for maintenance and leveling will be observed most strongly in individuals with extensive exposure to both a non-standard local variety and a regional or national standard.

Cynthia G. Clopper: Ohio State University. E-mail: clopper.1@osu.edu

1 Introduction

Synchronic variation is a hallmark characteristic of human speech, and arises from both intra-talker sources, such as speaking style and emotional state, and inter-talker sources, such as age, gender, and regional background. Ladefoged and Broadbent (1957) characterized this variation as arising both from physiological differences that convey personal information about who the talker is and from learned speech patterns that convey personal information, as well as socio-linguistic information about the groups to which the talker belongs. Klatt (1989)

similarly identified within- and between-talker sources of variability, such as the personal and sociolinguistic information identified by Ladefoged and Broadbent (1957), but also noted additional sources of variability, including recording (or listening) conditions and segmental variation due to linguistic factors such as coarticulation, articulatory undershoot, and phrasing.

Although speech perception and processing are generally robust to these sources of variability in good listening conditions (Pisoni 1997), listeners in more difficult listening conditions exhibit significant processing benefits for familiar voices and varieties relative to less familiar voices and varieties (see Pisoni [1993] for a review). The effects of talker familiarity on speech processing were tested explicitly in a series of studies by Nygaard and Pisoni (1998; Nygaard et al. 1994), in which they trained participants to identify a set of talkers by name. Following training, the participants exhibited better speech intelligibility in noise for those talkers whom they had learned to identify than for talkers whom they had not heard before. Similarly, Bradlow and Bent (2008) found that training participants to identify Chinese-accented speech led to improved intelligibility for unfamiliar Chinese-accented talkers, suggesting a talker-independent processing benefit for familiar accents.

The effects of talker familiarity have also been observed more indirectly in studies exploring the effects of talker variability on speech processing. For example, improvements in recognition memory have been observed for words repeated by the same talker relative to words repeated by a different talker (Craik and Kirsner 1974; Palmeri et al. 1993; Goldinger 1996). In addition, speech processing is slower and less accurate when the talker changes from trial to trial within an experiment than when the talker remains constant (Mullennix et al. 1989; Sommers et al. 1997). These findings suggest that even brief exposure to previously unfamiliar talkers within an experimental setting can provide listeners with sufficient familiarity with those talkers to facilitate processing. Rapid adaptation to variability has also been shown for foreign accents in as little as 2–4 utterances (Clarke and Garrett 2004).

These familiarity effects are consistent with models of speech processing, such as exemplar models, in which ease of processing reflects frequency of exposure to a particular word, talker, or accent (Johnson 1997; Pisoni 1997). In these models, more familiar exemplars are more strongly represented in the exemplar space than less frequently encountered exemplars and are therefore easier to process. In addition, recently encountered exemplars are more strongly represented in the exemplar space than exemplars with a longer time lag, which can account for the rapid adaptation to variation that is observed in the laboratory. The detailed information encoded in an exemplar model allows a single exemplar to simultaneously contribute to the representation of a number of different linguistic

and indexical categories. For example, my hearing a single utterance of the word *gas* produced by my neighbor will contribute to my representations of the phonemes /g, æ, s/, the word *gas*, and the categories for my neighbor's voice, adult female talkers, and native Ohioans, among others. These representations will facilitate my processing of future utterances of the phonemes /g, æ, s/ and the word *gas*, as well as future utterances produced by my neighbor and other adult female Ohioans. However, this simultaneous abstraction over multiple linguistic and indexical dimensions can lead to processing costs when utterances are highly variable in a task-irrelevant dimension (Mullennix et al. 1989; Mullennix and Pisoni 1990; Sommers et al. 1997). Thus, a recent exemplar of my neighbor saying the word *gas* may inhibit my processing of the word *guess* produced by a Southern male talker. Exemplar models in which multiple levels of linguistic and indexical information are simultaneously stored and represented can account for this complex pattern of observed facilitation and interference effects of indexical variation on linguistic processing.

Pierrehumbert (2002) extended this exemplar processing approach to include a model of speech production, in which production targets are randomly selected from the appropriate exemplar distribution. The distributions themselves are weighted according to the principles of frequency and recency that drive the perceptual models, so that production can shift over time within an individual as a result of both long-term experience and short-term novel exposure. This type of model, in which perception and production are intrinsically linked and both rely on accumulated exemplars, leads to several predictions about the effects of exposure to variation on speech processing and the potential for sound change. In particular, individuals with exposure to relatively little sociolinguistic variation should have relatively less variable distributions of phonological and/or lexical exemplars, whereas individuals with exposure to more sociolinguistic variation should have more variable distributions. These variable distributions are more likely to exhibit overlap, at the phonemic and/or lexical level, and language users' response to this overlap has the potential to result in the entrenchment of local forms and resistance to change under some conditions, but in the adoption of novel forms and dialect leveling under other conditions. As discussed in Section 2, behavioral evidence from speech perception tasks for these potential loci of change within the individual include processing benefits for the local variety, processing benefits for the standard variety, and both costs and benefits associated with exposure to multiple varieties. Whereas the processing benefits for the local variety may contribute to the entrenchment of local forms and maintenance of sociolinguistic variation, the processing benefits for the standard variety may contribute to change towards the standard and dialect leveling. The mixed costs and benefits observed in individuals with exposure to multiple varieties

suggest that the variable, experience-driven representations that are central to exemplar theory have competing potentials for maintenance vs. leveling of socio-linguistic variation.

2 Potential loci of change within the individual

2.1 Processing benefits for the local variety

Listeners exhibit processing benefits for their native regional dialect in a range of tasks, including speech intelligibility, lexical decision, and higher-level semantic processing. Mason (1946) provided the first evidence for a local dialect benefit in his intelligibility research for the U.S. Army. He found that Northerners and Southerners exhibited poorer intelligibility across regions than within regions in a standardized intelligibility test in noise. More recently, local dialect intelligibility benefits have been observed in phrase-final word recognition tests for listeners from Birmingham, Philadelphia, and Chicago (Labov and Ash 1997) and for listeners from different parts of Ohio (Flanigan and Norris 2000). In both studies, local listeners (from Birmingham and southern Ohio, respectively) more accurately identified the phrase-final target words than non-local listeners. Similar results have been obtained in isolated word recognition studies. For example, Preston (2005) reported higher overall accuracy for words containing Northern Cities shifted vowels for listeners from southern Michigan, where the shift is fairly advanced, than for listeners from rural, mid-Michigan, where the shift is less advanced. Jacewicz and Fox (2012) also reported overall local dialect benefits in a fully-crossed word recognition task comparing North Carolina and Wisconsin varieties.

With the exception of Mason's (1946) study, however, all of these intelligibility tests presented the materials in good listening conditions, which may not provide an ecologically sufficient test of cross-dialect intelligibility. Clopper and Tamati (2010), however, conducted a word recognition task in noise with talkers and listeners from the General American and Northern dialects of American English. The General American dialect corresponds to Labov's (1998) "Third Dialect" and includes the rhotic New England, Midland (southern midwestern), and Western varieties. The Northern dialect is spoken in the Great Lakes region in the northern midwestern United States, from upstate New York to Minnesota. Clopper and Tamati's (2010) experiment was designed to explore implicit recognition memory and therefore included two experimental blocks: a familiarization block consisting of 33 stimulus words in each dialect and a test block consisting of 66

Table 1: Word recognition accuracy in noise for General American and Northern talkers by General American and Northern listeners in the test block of Clopper and Tamati’s (2010) implicit recognition memory study.

	General American talkers	Northern talkers
General American listeners	68%	66%
Northern listeners	66%	71%

stimulus words (33 repeated and 33 new) in each dialect. In the familiarization block, an overall effect of talker dialect was observed, in which the Northern talkers were more intelligible than the General American talkers, consistent with our previous work on word recognition in noise with these materials (Clopper et al. 2010). As shown in Table 1, however, we found a significant local dialect benefit in the test block of the experiment, and an analysis of variance confirmed a statistically significant talker dialect $\times$ listener dialect interaction [$F(2, 47) = 5.3$, $p = .008$]. For the General American listeners ($N = 21$), word recognition was more accurate for the General American talkers than the Northern talkers [$t(20) = 2.3$, $p = .03$], whereas for the Northern listeners ($N = 16$), word recognition was more accurate for the Northern talkers than the General American talkers [$t(15) = -3.1$, $p = .008$].

Local dialect effects have also been observed in lexical decision tasks in French (Flocchia et al. 2006), Dutch (Impe et al. 2008), and American English (Sumner and Samuel 2009). In the French and Dutch studies, the local dialect benefit emerged in the participants’ relative response times: lexical decision judgments were faster for targets produced in the local dialect than for targets produced in a non-local dialect. In the American English study, the local dialect benefit emerged in the observed patterns of priming: local listeners exhibited greater form and semantic priming for the local variant than non-local listeners. In addition, Adank and her colleagues have shown that the local dialect benefit extends beyond semantic priming to other higher-level semantic processes, including animacy decisions (Adank and McQueen 2007) and sentence verification (Adank et al. 2009). Like the lexical decision judgments in French and Dutch, these kinds of semantic judgments are faster for utterances produced in the local dialect than a non-local dialect.

Taken together, these findings reveal that speech produced in the local, native variety is more intelligible and processed more quickly than speech produced in a less familiar regional dialect. This effect of dialect familiarity suggests a relatively strong cognitive representation of the native regional dialect. In an exemplar model, these strong representations could be characterized by tight,

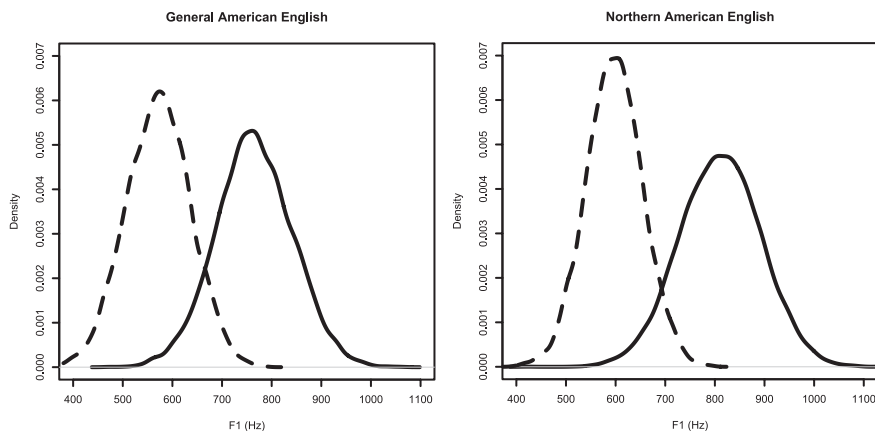


Fig. 1: Illustration of the low-variance, non-overlapping representations of two vowel categories (dashed, solid) in the native, local variety for a hypothetical General American listener (left) and a hypothetical Northern listener (right) with relatively little exposure to other sources of variation.

non-overlapping phonological or lexical distributions, in which the local forms are highly entrenched. To illustrate this pattern of exemplar distributions and facilitate comparison to the patterns of distributions discussed in the following sections, a series of random samples were generated and plotted in Figure 1. The density plots in Figure 1 provide a model of exposure to two vowel categories (/I, ε/) for hypothetical language users with exposure to only a single dialect, either the General American dialect (left) or the Northern dialect (right) of American English. The density plots were created by obtaining 10,000 random samples from each of four normal distributions defined by the observed means and standard deviations of the first formant frequencies extracted from 10 words containing the vowel /I/ and 10 words containing the vowel /ε/ from each of 10 female General American talkers and 10 female Northern talkers, respectively. Although each of these random samples was based on the distribution of real linguistic data to allow for a concrete interpretation of the patterns, the distributions that are modeled by these samples are an extreme simplification of human experience. The acoustic information is modeled in only a single dimension (F1) and the linguistic and indexical variation captured by the model variances is limited to 10 words produced in citation form by each of 10 female talkers. The exposure to a particular vowel category that an adult language user has accumulated is much greater in magnitude and complexity than these hypothetical samples, but estimation of true distribution parameters is intractable. Thus, these models should

be interpreted as simplified illustrations of the patterns of distributions that would lead to the observed processing results.

The hypothetical distributions for /ɪ/ (dashed line) and /ε/ (solid line) for the General American dialect are shown in the left panel of Figure 1, and the hypothetical distributions for the Northern dialect are shown in the right panel. Although the category means are higher for both vowels for the Northern dialect than the General American dialect, consistent with /ɪ, ε/ lowering in the Northern dialect (Labov 1998), the vowel category distributions show very little overlap within each dialect. Thus, for these two hypothetical language users, identifying a particular target utterance in their native dialect as either /ɪ/ or /ε/ should be relatively easy, even if the only information available is F1 frequency, because the low-variance, non-overlapping distributions suggest robust phonological categories. Robust categories are highly stable and unlikely to exhibit drift or change over time, unless the individual moves to an environment in which he or she encounters a large number of exemplars from a different social or regional distribution (see Sections 2.3 and 2.4). This entrenchment of local forms and resistance to change should contribute to the maintenance of sociolinguistic variation over time within and across individuals.

2.2 Processing benefits for the standard variety

Processing benefits are also observed for the ‘standard’ variety of a language, often represented in national media, regardless of a listener’s native variety. Similar to the local dialect benefit described in Section 2.1, the standard variety benefit is reflected in greater intelligibility and faster processing for the standard than for non-standard regional varieties. For example, standard General American English is more intelligible in noise than non-standard Southern or Mid-Atlantic American English for listeners from many different regions of the United States, including the Mid-Atlantic (Clopper and Bradlow 2008; Clopper 2012). Similarly, lexical decision responses are faster for words produced in the standard dialect of European French than a non-standard, non-local dialect (Floccia et al. 2006), and form and semantic priming are observed in lexical decision tasks for General American English, regardless of the native variety of the listener (Sumner and Samuel 2009). Faster lexical processing for General American English has also been observed in the speeded lexical classification task described by Clopper and Pate (2008). As shown in Table 2, lexical identification of target words as either *bad* or *bed* was significantly faster [$F(1, 23) = 70.62$, $p < .001$] and more accurate [$F(1, 23) = 27.94$, $p < .001$] for the General American targets than for the Northern targets over 64 speeded forced-choice trials. This

Table 2: Speeded lexical classification mean accuracy and response time for General American and Northern target words.

	Accuracy	Response time
General American targets	98%	552 ms
Northern targets	85%	643 ms

pattern of results was observed for both General American ($N = 16$) and Northern ($N = 9$) listeners.

Standard varieties also exhibit long-term form priming effects in lexical decision and word recognition tasks. Sumner and Samuel (2009) observed significant long-term form priming in their lexical decision task for General American prime-target pairs for both General American and Mid-Atlantic listeners. Clopper and Tamati (2010) similarly observed long-distance form priming for General American prime-target pairs in our implicit recognition memory task, but, as shown in Figure 2, we also observed significant form priming for all prime-target pairs in-

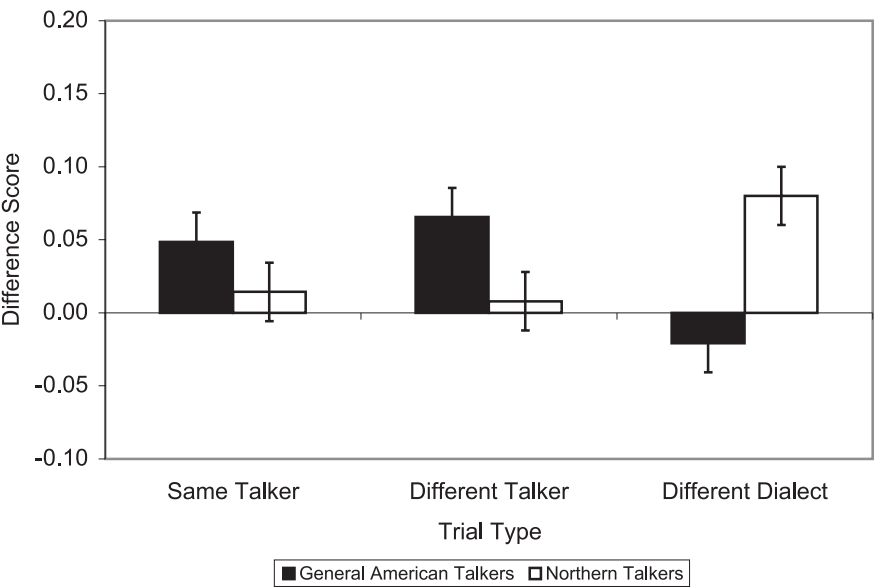


Fig. 2: Mean proportion correct difference scores for each trial type (Same Talker, Different Talker, and Different Dialect) for each talker dialect (General American and Northern) in Clopper and Tamati’s (2010) implicit recognition memory study. Error bars are standard error of subject means.

volving a General American prime. In particular, a significant repetition benefit, corresponding to an improvement in accuracy from the familiarization block to the test block for the repeated stimulus words, was observed for General American targets in the test block preceded in the familiarization block by an acoustically identical prime [Same Talker; $t(49) = 2.7$, $p = .009$], for General American targets in the test block preceded by a prime in the familiarization block produced by a different General American talker [Different Talker; $t(49) = 2.5$, $p = .016$], and for Northern targets in the test block preceded by a prime in the familiarization block produced by a General American talker [Different Dialect; $t(49) = 4.5$, $p < .001$]. However, no significant repetition benefits were observed for the Northern primes. Thus, regardless of the dialect of the target test word, a repetition or priming effect was observed only when the word was produced in the familiarization block by a General American talker. This effect was observed for both General American ($N = 21$) and Northern ($N = 16$) listeners.

Taken together, these results suggest that exposure to a standard variety, through education and/or the media, can lead to significant processing benefits for the standard variety, regardless of the listener's native dialect. The standard variety benefit can be accompanied by a local dialect benefit, so that lexical processing is fast and accurate for both the local and the standard varieties (e.g., Floccia et al. 2006; Sumner and Samuel 2009), but it may also occur at the expense of the local dialect, such that no local dialect benefit is observed (e.g., Clopper and Bradlow 2008; Clopper and Tamati 2010). This competition between the local and standard varieties may reflect recency effects, particularly in my work, which is conducted in a General American dialect region, or more general expectations about the kinds of speech one is likely to encounter in a university laboratory setting. In addition, although recent research by Stuart-Smith (2011; Stuart-Smith et al. 2013) suggests that media exposure can substantially impact language use, the effects of education and the media on exposure to the standard language variety are confounded for the studies discussed in this section because they were all conducted in a university setting with university student participants.

The standard variety benefit is consistent with a model in which frequent exposure to a particular variety, even passively through the media, can facilitate speech processing. In this account, for individuals whose native variety is not the standard, robust representations of the standard variety are stored along with the representations of the local variety. This pattern of distributions is illustrated in Figure 3. The left panel of Figure 3 shows the General American (gray) distributions of /ɪ, ε/ (dashed, solid) from Figure 1 overlaid on the Northern (black) distributions of /ɪ, ε/ (dashed, solid) from Figure 1, representing the vowel distributions of two hypothetical language users (gray and black) with exposure

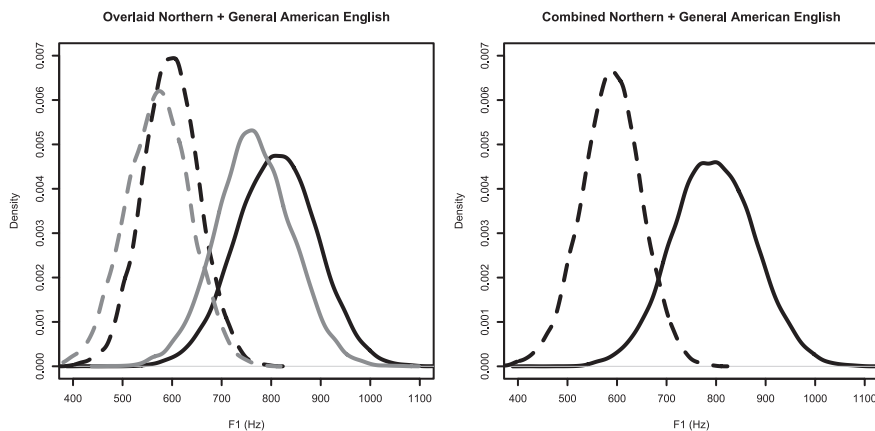


Fig. 3: Illustration of the difference in representation of two vowel categories (dashed, solid) for hypothetical listeners with exposure to one or both of the Northern and General American varieties. The left panel shows the distributions for the hypothetical single-dialect General American (gray) and Northern (black) listeners from Figure 1. The right panel shows the combined distributions for the hypothetical language user with exposure to both the Northern and General American varieties.

to different single varieties (General American or Northern). The right panel of Figure 3 shows the vowel distributions of a hypothetical language user with exposure to both the local Northern variety and the standard General American variety. These density plots were created by combining the 10,000 samples of the native Northern distributions with 5,000 random samples from the General American distributions. Fewer samples were selected from the General American dialect than the Northern dialect to model the hypothetical relative exposure to the standard and the local native varieties.

These combined distributions of the local and standard varieties in the right panel of Figure 3 are more complex than the distributions of the single varieties in the left panel of Figure 3. Note that the spread of both vowel distributions has increased from the Northern-only distributions in the left panel of Figure 3 to the combined Northern and General American distributions in the right panel of Figure 3. In addition, both vowel categories are beginning to show a small ‘shoulder’ on the left side of their distributions in the right panel, reflecting the leftward skewing influence of the standard General American distribution on the native Northern distribution. This kind of variable distribution may contribute to dialect leveling and the reduction of sociolinguistic variation over time as more exemplars from the standard variety are encountered and the means of the distributions shift towards the standard.

2.3 Processing benefits associated with exposure to multiple varieties

The preceding discussion of processing benefits associated with local and standard varieties suggests that listeners who have substantial exposure to multiple different dialects as a result of self-reported geographic mobility and/or exposure to a non-local standard through education or the media will exhibit a processing benefit for all of the varieties with which they are familiar. As noted above, this kind of predicted simultaneous benefit for the local and the standard varieties has been observed in lexical decision tasks, both with respect to overall response time (Floccia et al. 2006) and with respect to priming patterns (Sumner and Samuel 2009). In addition, Sumner and Samuel’s (2009) results reveal that familiarity with both the General American and Mid-Atlantic varieties facilitates cross-dialect priming. Although cross-dialect priming is weaker overall than within-dialect priming, local listeners, who have substantial exposure to both the local and the standard varieties, exhibit cross-dialect form and semantic priming in both directions: General American forms prime Mid-Atlantic forms and vice versa. However, non-local listeners, who have much less experience with the Mid-Atlantic variety, exhibit much more selective cross-dialect priming.

Processing benefits are also observed in cross-dialect intelligibility tasks for individuals who have extensive experience with multiple dialects, particularly when two sounds are phonetically merged or nearly merged in one of the two varieties. For example, as shown in Table 3, Clopper et al. (2010) found that Northern listeners ($N = 9$) were less likely to confuse /ɔ/ with /ɑ/ than General American listeners ($N = 21$) for both Northern and General American targets. This finding is consistent with the observation that the /ɔ, ɑ/ merger is incomplete throughout much of the General American dialect region (Labov et al. 2005): although many General American listeners perceive /ɔ, ɑ/ to be homophones, their productions may remain acoustically distinct. Thus, as a result of their experience with both the merging standard variety and their own unmerged native

Table 3: Word recognition mean accuracy for /ɔ/ for General American and Northern targets by General American and Northern listeners in Clopper et al.’s (2010) cross-dialect word recognition in noise study.

	General American talkers	Northern talkers
General American listeners	66%	69%
Northern listeners	76%	87%

variety, the Northern listeners in Clopper et al.'s (2010) study were able to use the small acoustic differences between the vowels produced by the General American talkers to accurately identify them. Similar results would be expected for native listeners of other unmerged non-standard varieties, including Mid-Atlantic and Southern American English.

This result differs from Dufour et al.'s (2007) finding that mere exposure to a contrast in the standard variety that is not present in the local variety can lead to discrimination between two phonemes in a lexical decision task. In their study, listeners from southern France, for whom the contrast between /o, ɔ/ is merged, exhibited the same lack of semantic priming for /o, ɔ/ minimal pairs as listeners from Geneva, who maintain the contrast. Dufour et al. (2007) interpreted this result as revealing the maintenance of the contrast in perception for the southern French listeners, despite the lack of a contrast in production. This difference between the French and American data may reflect explicit social stereotypes about the relevant mergers. Dufour et al. (2007) noted that southern French listeners have explicit knowledge about the /o, ɔ/ contrast in standard French, whereas the merger of /ɔ, a/ in American English is occurring in non-stereotyped varieties and is not subject to overt commentary.

Finally, Warren and his colleagues (Rae and Warren 2002; Warren and Hay 2006) have shown that experience with a merger-in-progress can also shape semantic priming patterns. In New Zealand, the vowels /i/ and /e/ are merging as [i] before /r/ (which is realized as [ə] in this non-rhotic variety). Semantic priming patterns reflect the asymmetry of this merger: although [ʃiə] primes both *shout* and *sit*, because it is a plausible phonetic realization of both *cheer* and *chair*, [ʃeə] primes only *sit*. Crucially, listeners' exposure to multiple stages of the change-in-progress and a cognitive representation of the variability associated with the ongoing merger are essential for accounting for these results.

Taken together, these findings are consistent with a model in which exposure to multiple different varieties leads to more variable representations. These variable representations are defined by distributions with greater variance or bandwidth than the distributions of less variable input, which provides a larger space of potential mappings between the signal and linguistic categories. The cross-dialect processing benefits observed for participants with exposure to multiple varieties reflect this greater flexibility in perceptual mapping, because this flexibility provides access to both local and standard forms in speech processing tasks. That is, a given stimulus can be mapped onto more than one linguistic representation, which increases the likelihood that the stimulus will be accurately identified. However, this one-to-many mapping can also lead to processing costs due to competition between possible linguistic categories. These costs will be discussed in Section 2.4.

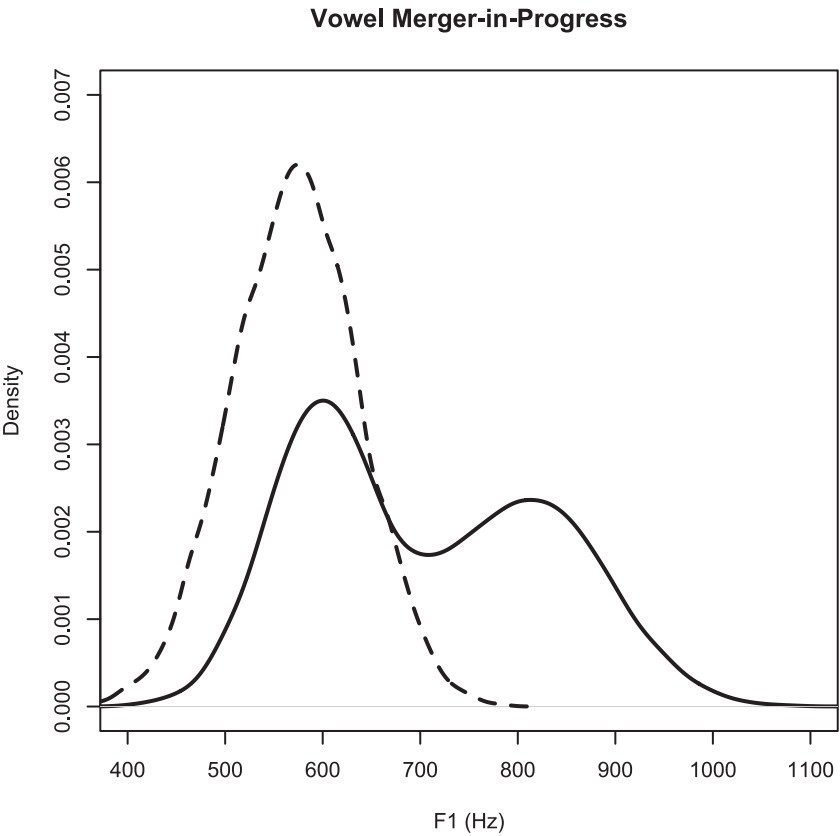


Fig. 4: Illustration of the representation of a merger-in-progress of two vowel categories. The solid line represents the bimodal distribution of the merging vowel category. The dashed line represents the distribution of the vowel category with which the solid category is merging.

The variable representations associated with a hypothetical merger-in-progress of /ε/ with /ɪ/ in American English are illustrated in Figure 4. The Northern samples of /ɪ, ε/ were combined to obtain a bimodal distribution for /ε/ (solid line). This bimodal distribution reflects the merger of /ε/ with /ɪ/ over time across the speech community. The higher F1 values for /ε/ correspond to older speakers who maintain the /ɪ, ε/ distinction, whereas the lower F1 values correspond to younger speakers who exhibit the merger. The General American samples of /ɪ/ (dashed line) represent the stable distribution of /ɪ/ across the speech community. The resulting pattern is two partially overlapping distributions in which vowels with an F1 value of 600 Hz can correspond to either vowel category, but vowels with an F1 of 800 Hz most likely correspond to /ε/. This kind of

representation predicts the results obtained by Warren and his colleagues for the New Zealand /i, e/ merger before /r/: an utterance with an unambiguous [e] is only consistent with the /e/ distribution, whereas an utterance with an [i] is consistent with either interpretation because the distributions exhibit substantial overlap in that part of the phonetic space.

2.4 Processing costs associated with exposure to multiple varieties

Overlapping distributions and one-to-many mappings can also result in processing costs. Whereas standard and stereotyped local variants can facilitate processing for listeners who are familiar with both the standard and the local forms, local variants that are not stereotyped may inhibit processing for listeners who have experience with both forms. For example, although Dufour et al. (2007) found similar patterns of priming for /o, ɔ/ minimal pairs for the southern French and Geneva listeners, they did not observe similar priming patterns for /e, ε/ minimal pairs. In the latter case, the southern French listeners reliably perceived /e, ε/ as homophones, despite their contrastive status in the standard variety.

Similarly, Clopper (2011) observed inhibition for competing /ε, æ/ minimal pair primes for Northern American English listeners in a cross-dialect lexical decision task. When Northern listeners ($N = 24$) were presented with an auditory prime such as *bad*, they were slower to identify its minimal pair *bed* as a real word than to identify an unrelated word, such as *knit*, as a real word. As shown in the left panel of Figure 5, this inhibition was observed for both General American and Northern talker dialects, but only for Northern listeners ($t(23) = 3.27$, $p = .003$). The General American listeners ($N = 42$) did not exhibit significant inhibition. This effect was partially replicated by Clopper and Walker (2013), as shown in the right panel of Figure 5. We found inhibition due to minimal pair competition for all listener groups, but marginally more inhibition for Mobile listeners ($N = 50$), who have lived in multiple different dialect regions and therefore have extensive exposure to multiple dialects, than for General American listeners ($N = 56$), who have lived exclusively in the General American region and therefore have more limited experience with variation ($F(1, 104) = 3.82$, $p = .053$).

Finally, larger cross-dialect interference effects are observed in speeded lexical classification for Northern listeners, who are highly familiar with both the standard and their own non-standard regional variety, than for General American listeners, who are less familiar with the non-standard variety. In the task described by Clopper and Pate (2008), the target words *bad* and *bed* were presented

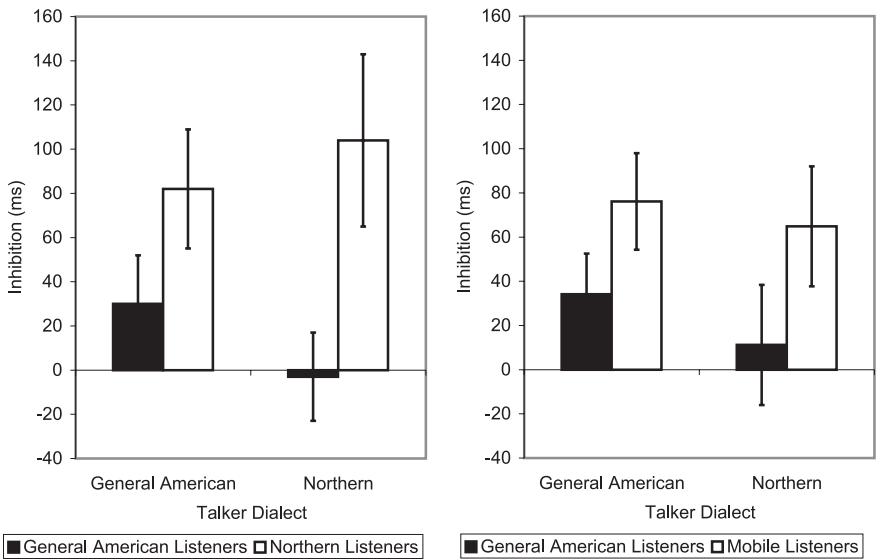


Fig. 5: Inhibition due to minimal pair competition for General American and Northern talkers by General American, Northern (left panel only), and Mobile (right panel only) listeners in Clopper's (2011; left) and Clopper and Walker's (2013; right) lexical decision studies.

to listeners for identification in single-dialect (General American or Northern) blocks of 16 trials each and in a mixed-dialect (General American and Northern) block of 32 trials. Although both General American ($N = 16$) and Northern ($N = 9$) listeners exhibited slower responses in the mixed-dialect block than the single-dialect blocks ($F(1, 23) = 50.62, p < .001$), this dialect interference effect was significantly larger for the Northern listeners than the General American listeners ($F(1, 23) = 4.34, p = .049$), as shown in Table 4.

Taken together, these results suggest that exposure to multiple dialects can result in processing costs when phonetic variability across the familiar dialects leads to greater competition among lexical representations. As discussed in

Table 4: Mean response times and dialect interference effect for General American and Northern listeners in the speeded lexical classification task.

	General American listeners	Northern listeners
Single-dialect block	546 ms	534 ms
Mixed-dialect block	641 ms	646 ms
Interference effect	95 ms	112 ms

Section 2.3, exposure to multiple dialects leads to more variable representations, such as those illustrated in the right panel of Figure 3 and in Figure 4, and those variable representations provide a larger space of potential mappings between the signal and linguistic categories. The interference effects observed in these speeded lexical processing tasks for listeners with frequent exposure to multiple varieties suggest competition in these perceptual mappings. Thus, although in the experiments discussed in Section 2.3, the one-to-many mapping led to greater flexibility in speech processing, in the experiments discussed here, the one-to-many mapping led to increased lexical competition and slower processing. The experiments described in these two sections used a number of different tasks to explore a range of linguistic phenomena that exhibit varying degrees of stereotyping. Given the documented effects of social and linguistic stereotypes on speech processing (Rubin and Smith 1990; Niedzielski 1999; Strand 1999; Hay and Drager 2010), further research is essential to disentangle the effects of these factors on facilitation vs. interference of multi-dialect exposure in speech processing.

In an exemplar model, the competing linguistic representations would be realized with greater variance and potentially bimodal distributions of exemplars. These kinds of distributions are inherently less stable than the kind of well-separated unimodal distributions illustrated in Figure 1 and may therefore be more susceptible to drifting and changing over time. Drift towards the local variety will result in maintenance of sociolinguistic variation, whereas drift towards the standard will result in dialect leveling. Thus, individuals with exposure to multiple dialects have the potential to drive both sociolinguistic change and maintenance, depending on the direction of the drift in their representations over time.

3 Towards a model of change within the individual

Taken together, these effects of familiarity on cross-dialect speech processing suggest that long-term exposure to synchronic variation affects the representation and processing of linguistic information. The processing benefits associated with familiar varieties suggest that language users have strong cognitive representations for these varieties. However, the interference effects observed for listeners with frequent exposure to multiple dialects suggest that the maintenance of multiple competing linguistic systems can negatively impact speech processing. The competition between strong representations of a native regional variety and the standard variety may provide a locus of sound change for these indi-

viduals because, although the representations of local forms are stable and entrenched, substantial experience with a second variety may lead to change as the mean of the phonological or lexical distribution drifts as a function of accumulating experience with the second variety.

Crucially, an exemplar-based perspective allows us to account for both maintenance of sociolinguistic variation through the entrenchment of local forms and for dialect leveling through change towards the standard. In particular, maintenance can be achieved when the exemplars associated with the local variety are more frequent and/or more heavily weighted in processing than the standard variety. In addition to maintenance, local variants may change away from the standard if the combined distributions become too large and overlapping. As in second language acquisition, the variant in one system may be “deflected” away from the variant in the other system to preserve the phonetic distinction between the two variants (and varieties) within the phonological system (Flege 1995). Although in second language acquisition this process is predicted to result in non-native production of the second language category, divergence from the standard is a well-established phenomenon in sociolinguistic variation, and a local language community can have a much larger impact on its own variety than on the standard variety that is shared by a much larger speech community. Thus, competition between local and standard forms could lead to divergence of the local variety away from the standard variety. The alternative to maintenance or divergence is dialect leveling, in which increased experience with and relatively heavier weighting of the standard variants may lead to a change towards the standard. Thus, maintenance, divergence, and dialect leveling are all potential outcomes of within-talker cross-dialect exposure.

The primary factors determining which outcome emerges remain to be determined empirically, but research in the psycholinguistic and sociolinguistic traditions suggests that frequency of exposure and social status and solidarity among interlocutors will play important roles in the structure and weighting of these variable representations. For example, in the psycholinguistic literature on accommodation, Goldinger (1998) showed that low-frequency words with a relatively high frequency of experimental exposure exhibit the greatest degree of imitation or accommodation in a word repetition task, suggesting that the combination of long-term experience (lexical frequency) and short-term exposure (experimental repetitions) significantly affects production targets, consistent with Pierrehumbert’s (2002) proposed exemplar model integrating speech perception and production. Research in social psychology and sociolinguistics, however, has shown that social factors, including the negotiation of power and solidarity in an interaction, can also affect the degree to which interlocutors converge or diverge with one another (e.g., Giles et al. 1987; Coupland and Giles 1988). Social factors

have also recently been shown to affect imitation in the laboratory (Pardo 2006; Babel 2012). These results suggest that talkers have some ability to (explicitly or implicitly) select production targets that will help them achieve their communicative goals.

In an exemplar model, these social factors would need to affect the relative weights assigned to stored exemplars, so that production targets are randomly selected from the subset of exemplars that reflect the social variants that the talker wishes to produce. Thus, convergence would emerge when stronger weights are assigned to the recent exemplars of the interlocutor, whereas divergence would emerge when stronger weights are assigned to some other cohort of exemplars, such as those associated with a particular social category. This kind of reweighting of exemplars is consistent with observations about adaptation and generalization in perception. Rapid adaptation to variation has been demonstrated not only for the perception of foreign accents (e.g., Clarke and Garrett 2004), but also in research exploring variation that is plausibly sociolinguistic. For example, Dahan et al. (2008) and Maye et al. (2008) both observed perceptual adaptation to vowel shifts after very brief exposure in the laboratory. Kraljic et al. (2008) observed similar rapid adaptation to contextually-driven consonant variability. The results of these studies suggest that listeners require very little experience with a particular talker to retune their perceptual processes to accommodate potential sociolinguistic variation.

However, Kraljic and Samuel (2011) found that this accommodation can be blocked if there is external information about an irrelevant source of the variation. In particular, if the non-standard variant can be attributed to the talker having a pen in his or her mouth, listeners will not adjust their perceptual categories. This evidence that both signal-internal and signal-external information about social variation can affect perceptual adjustment is consistent with the view that language users have some control over the relative weighting of exemplars in the perception-production space. When variation can be attributed to talker-specific or sociolinguistic factors, listeners rely on previous experience with that variation in future encounters with that talker, but when variation is attributable to external, temporary factors, listeners do not rely on those deviant exemplars in future processing. Similarly, to create social solidarity or distance, talkers may be able to rely relatively more or less, respectively, on the recent exemplars of their interlocutor in choosing their own production targets. Over time, these individual adjustments may lead to significant shifts in phonological and/or lexical distributions that correspond to community-level dialect maintenance or leveling. Garrett and Johnson (2012) presented some preliminary simulations of the effects of social identity on language change that are consistent with this model of exemplar weighting.

Current theories of community-level variation and change are also consistent with this perspective on the locus of sound change in individuals and their relative weighting of exemplars. For example, women are often described as the leaders of linguistic change, but they are also observed to produce fewer stigmatized variants of stable social variables than men (Labov 2001). Eckert (1989) attributed this pattern of variation to women's greater use of linguistic variation to signal power and group affiliation than men. Similarly, Milroy and Milroy (1993) argued that individuals with weak social networks are more likely to be linguistic innovators than individuals with strong networks, because individuals with weak social group connections are more susceptible to outside influences than individuals in close-knit groups. Both of these patterns can be explained from the perspective of exemplar weighting. Women assign relatively more weight to new and socially prestigious forms than men, which leads to shifts in their distributions over time towards innovative and standard forms, respectively. Individuals with weak network ties assign relatively less weight to in-group variants, which leads to stronger influences of out-group variants on the shape of their overall distributions over time. This change within an individual as a result of the relative weighting of exemplars crucially requires the language user to encounter sufficient variation, such as between a local variety and a standard variety, to allow the distribution to change over time.

Although a complete understanding of the relationships among perceptual adaptation in the individual, production variability in the individual, and sound change in the community awaits a deeper integration of the work presented here with work in accommodation, second language acquisition, and language variation and change, the experimental results described in Section 2 demonstrate the significant role of experience in shaping linguistic representations, and the effect of those complex representations on cross-dialect speech processing. I have further argued that the observed effects of experience are consistent with an exemplar-based model of linguistic representation, such as that proposed by Pierrehumbert (2002), and that understanding how synchronic variation is cognitively represented has implications for models of sound change, particularly at the level of the individual.

Acknowledgments: The research described in this paper was conducted in collaboration with Ann Bradlow, John Pate, Janet Pierrehumbert, Terrin Tamati, and Abby Walker. I would also like to thank Lauren Eskin, Joselyn Gilbert, Sarah Glazier-Kipe, Kenney Hensley, Erin Loxley, Jane Mitsch, and Haley Schram for assistance with data collection for these various projects. Portions of this work were supported by NIH F32 DC007237 to Northwestern University and NSF BCS-1056409 to Ohio State University.

References

- Adank, Patti, Bronwen G. Evans, Jane Stuart-Smith, & Sophie K. Scott. 2009. Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance* 35. 520–529.
- Adank, Patti, & James M. McQueen. 2007. The effect of an unfamiliar regional accent on spoken word comprehension. *Proceedings of the 16th International Congress of Phonetic Sciences*. 1925–1928.
- Babel, Molly. 2012. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics* 40. 177–189.
- Bradlow, Ann R., & Tessa Bent. 2008. Perceptual adaptation to non-native speech. *Cognition* 106. 707–729.
- Clarke, Constance M., & Merrill F. Garrett. 2004. Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America* 116. 3647–3658.
- Clopper, Cynthia G. 2011. Effects of naturally-occurring segmental ambiguity on cross-modal priming of minimal pairs. *Journal of the Acoustical Society of America* 129. 2659.
- Clopper, Cynthia G. 2012. Effects of dialect variation on the semantic predictability benefit. *Language and Cognitive Processes* 27. 1002–1020.
- Clopper, Cynthia G., & Ann R. Bradlow. 2008. Perception of dialect variation in noise: Intelligibility and classification. *Language and Speech* 51. 175–198.
- Clopper, Cynthia G., & John K. Pate. 2008. Effects of talker and token variability on perceptual learning of dialect categories. *Proceedings of Meetings on Acoustics* 5(060002).
- Clopper, Cynthia G., Janet B. Pierrehumbert, & Terrin N. Tamati. 2010. Lexical neighborhoods and phonological confusability in cross-dialect word recognition in noise. *Laboratory Phonology* 1. 65–92.
- Clopper, Cynthia G., & Terrin N. Tamati. 2010. Lexical recognition memory across dialects. *Journal of the Acoustical Society of America* 127. 1956.
- Clopper, Cynthia G., & Abby Walker. 2013. Effects of linguistic experience and phonetic ambiguity on lexical competition. Paper presented at Experimental Approaches to Perception and Production of Language Variation, University of Copenhagen, 20–22 March.
- Coupland, Nikolas, & Howard Giles. 1988. Introduction: The communicative contexts of accommodation. *Language and Communication* 8. 175–182.
- Craik, Fergus I. M., & Kim Kirsner. 1974. The effect of speaker's voice on word recognition. *Quarterly Journal of Experimental Psychology* 26. 274–284.
- Dahan, Delphine, Sarah J. Drucker, & Rebecca A. Scarborough. 2008. Talker adaptation in speech perception: Adjusting the signal or the representations? *Cognition* 108. 710–718.
- Dufour, Sophie, Noël Nguyen, & Ulrich Hans Frauenfelder. 2007. The perception of phonemic contrasts in a non-native dialect. *JASA Express Letters* 121. EL131–EL136.
- Eckert, Penelope. 1989. The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1. 245–267.
- Flanigan, Beverly Olson, & Franklin Paul Norris. 2000. Cross-dialectal comprehension as evidence for boundary mapping: Perceptions of the speech of southeastern Ohio. *Language Variation and Change* 12. 175–201.
- Fllege, James Emil. 1995. Second-language speech learning: Theory, findings, and problems. In Winifred Strange (ed.), *Speech perception and linguistic experience*, 233–277. Timonium, MD: York Press.

- Floccia, Caroline, Jeremy Goslin, Frédérique Girard, & Gabrielle Konopczynski. 2006. Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance* 32. 1276–1293.
- Garrett, Andrew, & Keith Johnson. 2012. Phonetic bias in sound change. In Alan C. L. Yu (ed.), *Origins of sound change: Approaches to phonologization*, 51–97. Oxford: Oxford University Press.
- Giles, Howard, Anthony Mulac, James J. Bradac, & Patricia Johnson. 1987. Speech Accommodation Theory: The first decade and beyond. *Communication Yearbook* 10. 13–48.
- Goldinger, Stephen D. 1996. Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22. 1166–1183.
- Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105. 251–279.
- Hay, Jennifer, & Katie Drager. 2010. Stuffed toys and speech perception. *Linguistics* 48. 865–892.
- Impe, Leen, Dirk Geeraerts, & Dirk Speelman. 2008. Mutual intelligibility of standard and regional Dutch language varieties. *International Journal of Humanities and Arts Computing* 2. 101–117.
- Jacewicz, Ewa, & Robert Allen Fox. 2012. The effects of cross-generational and cross-dialectal variation on vowel identification and classification. *Journal of the Acoustical Society of America* 131. 1413–1433.
- Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In Keith Johnson & John W. Mullenix (eds.), *Talker variability in speech processing*, 145–165. San Diego: Academic Press.
- Klatt, Dennis H. 1989. Review of selected models of speech perception. In William Marslen-Wilson (ed.), *Lexical representation and process*, 169–226. Cambridge, MA: MIT Press.
- Kraljic, Tanya, Susan E. Brennan, & Arthur G. Samuel. 2008. Accommodating variation: Dialects, idiolects, and speech processing. *Cognition* 107. 54–81.
- Kraljic, Tanya, & Arthur G. Samuel. 2011. Perceptual learning evidence for contextually-specific representations. *Cognition* 121. 459–465.
- Labov, William. 1998. The three dialects of English. In Michael D. Linn (ed.), *Handbook of dialects and language variation*, 39–81. San Diego: Academic Press.
- Labov, William. 2001. *Principles of linguistic change: Social factors*. Malden, MA: Blackwell.
- Labov, William, & Sharon Ash. 1997. Understanding Birmingham. In Cynthia Bernstein, Thomas Nunnally, & Robin Sabino (eds.), *Language variety in the south revisited*, 508–573. Tuscaloosa: University of Alabama Press.
- Labov, William, Sharon Ash, & Charles Boberg. 2005. *Atlas of North American English*. New York: Mouton de Gruyter.
- Ladefoged, Peter, & D. E. Broadbent. 1957. Information conveyed by vowels. *Journal of the Acoustical Society of America* 29. 98–104.
- Mason, Harry M. 1946. Understandability of speech in noise as affected by region of origin of speaker and listener. *Speech Monographs* 13(2). 54–68.
- Maye, Jessica, Richard N. Aslin, & Michael K. Tanenhaus. 2008. The Weckud Wetch of the Wast: Lexical adaptation to a novel accent. *Cognitive Science* 32. 543–562.
- Milroy, James, & Lesley Milroy. 1993. Mechanisms of change in urban dialects: The role of class, social network and gender. *International Journal of Applied Linguistics* 3. 57–77.

- Mullennix, John W., & David B. Pisoni. 1990. Stimulus variability and processing dependencies in speech perception. *Perception and Psychophysics* 47. 379–390.
- Mullennix, John W., David B. Pisoni, & Christopher S. Martin. 1989. Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America* 85. 365–378.
- Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18. 62–85.
- Nygaard, Lynne C., & David B. Pisoni. 1998. Talker-specific learning in speech perception. *Perception and Psychophysics* 60. 355–376.
- Nygaard, Lynne C., Mitchell S. Sommers, & David B. Pisoni. 1994. Speech perception as a talker-contingent process. *Psychological Science* 5. 42–46.
- Palmeri, Thomas J., Stephen D. Goldinger, & David B. Pisoni. 1993. Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19. 309–328.
- Pardo, Jennifer S. 2006. On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America* 119. 2382–2393.
- Pierrehumbert, Janet B. 2002. Word-specific phonetics. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology* 7, 101–139. Berlin: Mouton de Gruyter.
- Pisoni, David B. 1993. Long-term memory in speech perception: Some new findings on talker variability, speaking rate and perceptual learning. *Speech Communication* 13. 109–125.
- Pisoni, David B. 1997. Some thoughts on “normalization” in speech perception. In Keith Johnson & John W. Mullennix (eds.), *Talker variability in speech processing*, 9–32. San Diego: Academic Press.
- Preston, Dennis R. 2005. Belle’s body just caught the fit gnat: The perception of Northern Cities shifted vowels by local speakers. *Penn Working Papers in Linguistics* 11(2). 133–146.
- Rae, Megan, & Paul Warren. 2002. Goldilocks and the three beers: Sound merger and word recognition in NZE. *New Zealand English Journal* 16. 33–41.
- Rubin, Donald L., & Kim A. Smith. 1990. Effects of accent, ethnicity, and lecture topic on undergraduates’ perceptions of nonnative English-speaking teaching assistants. *International Journal of Intercultural Relations* 14. 337–353.
- Sommers, Mitchell S., Karen Iler Kirk, & David B. Pisoni. 1997. Some considerations in evaluating spoken word recognition by normal-hearing, noise-masked normal-hearing, and cochlear implant listeners I: The effects of response format. *Ear and Hearing* 18. 89–99.
- Strand, Elizabeth A. 1999. Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology* 18. 86–99.
- Stuart-Smith, Jane. 2011. The view from the couch: Changing perspectives on the role of the television in changing language ideologies and use. In Tore Kristiansen & Nikolas Coupland (eds.), *Standard languages and language standards in a changing Europe*, 223–239. Oslo: Novus Press.
- Stuart-Smith, Jane, Gwilym Pryce, Claire Timmins, & Barrie Gunter. 2013. Television can also be a factor in language change: Evidence from an urban dialect. *Language* 89. 501–536.
- Sumner, Meghan, & Arthur G. Samuel. 2009. The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language* 60. 487–501.
- Warren, Paul, & Jen Hay. 2006. Using sound change to explore the mental lexicon. In C. Fletcher-Flinn & G. Haberman (eds.), *Cognition and language: Perspectives from New Zealand*, 105–125. Bowen Hills, Queensland: Australian Academic Press.